

WHAT IS CLAIMED IS:

1. A traffic monitor for use with a set of one or more Web servers for providing statistical analyses of traffic, comprising:
 - an activity input for receiving data related to events on the set of servers;
 - means for categorizing events into categories;
 - means for associating events with subjects, wherein counts are maintained for each subject and subjects are associated with categories;
 - a normalizer for normalizing counts for events over a field of events; and
 - a result output for outputting results of the normalizer as the statistical analyses of traffic.

2. The traffic monitor of claim 1, wherein the activity input is an input from Web server logs.

3. The traffic monitor of claim 1, wherein the events include indications of page views, indications of search terms and indications of click streams of visitors to the set of servers.

4. The traffic monitor of claim 1, wherein the field of events is all page views.

5. The traffic monitor of claim 1, wherein the field of events is all page views in one category and wherein the normalizer normalizes a count for events over the single category field of events.

6. The traffic monitor of claim 1 comprising:
a canonicalization table that relates terms that can be represented by a canonized form;
a canonicalizer for generating at least one canonized term for an input term indicative of the input event;
a categorizer that generates one or more category indications for an input event; and
logic to assign a category to the input event based on the at least one canonized term generated by the canonicalizer for the input term indicative of the input event.

7. The traffic monitor of claim 1 comprising:

2 a click stream input that provides indications of navigation of a user subsequent to an
3 event; and
4 a categorizer that generates one or more category indications for an input event; and
5 logic to assign a category to the input event based on the indications of navigation of
6 a user subsequent to the event.

8. A method of generating statistics about traffic between a set of servers and a set of clients, comprising:

reading a log of events, wherein an event is a result of a client of the set of clients

making a request of a server of the set of servers and the server providing a response to the client;

automatically associating each event with one or more subject, wherein a subject is a topic or a term;

determining if a subject for an event is a canonical equivalent of another subject;

identifying one or more category relevant to the subject;

accumulating counts for events by subject, wherein counts for canonical equivalents

are accumulated together; and

outputting the accumulated counts.

9. The method of claim 8, wherein the set of servers is a constrained set of servers.

10. The method of claim 9, wherein the constrained set of servers comprises the servers for a portal Web site.

11. The method of claim 9, wherein the constrained set of servers comprises the servers for a plurality of portal Web sites.

12. The method of claim 8, wherein the set of servers is one server.

13. The method of claim 8, wherein the set of clients is an unconstrained set of clients.

14. The method of claim 8, wherein the set of clients is a constrained set of clients.

27. The method of claim 8, wherein the log of events includes a log of purchase transactions.

28. The method of claim 8, wherein the log of events includes a log of downloaded media objects.

29. The method of claim 8, further comprising a step of normalizing counts for each subject in a category relative to counts over the category.

30. The method of claim 8, wherein the step of associating an event with a subject, wherein the event is a search request, comprises the steps of:

- providing the client with search results responsive to the search request;
- recording a selection made by the client from the search results; and
- associating the search request with the subject of the selection.

31. The method of claim 8, further comprising the steps of:
determining a set of one or more demographic parameters relating to clients making requests or the users using the clients; and
using the determined set of one or more demographic parameters to partition the counts by demographic divisions.

32. The method of claim 8, further comprising the steps of:
determining a set of one or more demographic parameters relating to clients making requests or the users using the clients; and
using the determined set of one or more demographic parameters to determine a distribution of at least one count for a topic or term over a plurality of demographic divisions.

33. The method of claim 8, further comprising a step of generating a report showing comparisons of the traffic for each of a plurality of subjects in one or more categories.

34. The method of claim 8, further comprising a step of allocating advertising space based on the accumulated counts.

35. The method of claim 8, further comprising the steps of:

2 collecting traffic data prior to a campaign;
3 executing the campaign;
4 collecting traffic data after the campaign; and
5 comparing the traffic before and after the campaign as a measure of campaign
6 effectiveness.

1 36. The method of claim 35, wherein the campaign is a political
2 campaign, a marketing campaign, a general awareness campaign, a public service
3 announcement campaign, or a combination thereof.

1 37. The method of claim 8, further comprising a step of performing
2 intersection analysis.

1 38. The method of claim 8, further comprising a step of performing
2 associated interests analysis.

1 39. The method of claim 8, further comprising a step of generating an
2 advertisement wherein content of the advertisement is a function of the traffic statistics.

1 40. A method of accumulating counts for categories and subjects of search
2 events, comprising the steps of:
3 receiving, as a server, a search request from a client;
4 searching a set of items using search parameters of the search request;
5 providing the client with search results comprising a subset of the set of items
6 wherein the items in the subset have a predefined search criteria relationship to
7 the search parameters;
8 accepting a selection from the user of one of the subset of items; and
9 accumulating a count for the search event as a count for a subject or category
10 associated with a subject or category of the selection.

1 41. A method of canonicalizing search terms, comprising the steps of:
2 determining a first frequency of occurrence of a search term over a first period;
3 determining a second frequency of occurrence of a search term over a second period,
4 wherein the first period is prior to the second period;

5 if an increase in frequency from the first frequency to the second frequency is not
6 above a predetermined threshold, performing a first canonicalization process on
7 the search term;
8 if the increase in frequency is above the predetermined threshold, performing a
9 second canonicalization process on the search term, where the second
10 canonicalization process is more aggressive than the first canonicalization
11 process.

42. A method of canonicalizing search terms, comprising the steps of:
determining a first frequency of occurrence of a first search term over time;
determining a second frequency of occurrence of a second search term over time,
wherein the second search term is potentially canonically equivalent to the first
search term;
if the first frequency and the second frequency rise together, associating the first
search term and the second search term as canonical equivalents; and
if the first frequency and the second frequency do not rise together, not associating
the first search term and the second search term as canonical equivalents.